Available online at www.sciencedirect.com
SCIENCE DIRECT•
Applied Mathematics Letters 17 (2004) 1397–1401
Applied Mathematics Letters
www.elsevier.com/locate/aml

An Alternative Approach to Understanding Second-Order Conditions for Constrained Optimization—Lagrange Multipliers

E. D. MABERLY
Professor of Finance, University of Canterbury
Christchurch, New Zealand
edwin.maberly@canterbury.ac.nz
R. M. PIERCE
Senior Lecturer, Lincoln University
Canterbury, New Zealand

(Received and accepted July 2003)

Communicated by Graeme Wake

Abstract—This paper presents an alternative approach to solving a standard problem, frequently encountered in advanced microeconomics, using the technique of Lagrange multipliers. The objective is to enhance the understanding of students as to the derivation of the second-order conditions.

Keywords—Constrained optimization, Lagrange multipliers, Second-order conditions.

1. INTRODUCTION

Understanding the theoretical underpinnings of academic disciplines, like economics and finance, requires knowledge of mathematics that students often find difficult to master. Solutions that have a geometric or real world interpretation are helpful in closing the gap between a student's ability to apply an equation versus understanding its derivation.

In this paper, the second-order conditions for maximization of a function of two variables, subject to a linear constraint, are derived using an alternative methodology. The motivation is to provide a more intuitive approach to the "utility maximization problem", but, in particular, derivation of the second-order conditions. The utility maximization problem, which is described below, is typically solved by the use of Lagrange multipliers and forms an integral part of most, if not all, advanced microeconomics papers.

All students in advanced microeconomics are shown the technique of Lagrange multipliers for deriving necessary and sufficient conditions for maximizing the utility function, $U = f(x, y)$, subject to the budget constraint, $\bar{M} = \bar{P}_x x + \bar{P}_y y$. For simplicity, assume a two-product world, Product A and Product B, with the quantity consumed of each product, represented by the

0893-9659/04/$ - see front matter
doi:10.1016/j.aml.2003.07.012
Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

variables x and y with $\bar{P}_x$ and $\bar{P}_y$, their respective unit prices. The initial endowment, $\bar{M}$, is spent on the two products. The utility function $U = f(x,y)$ is unique to each individual, as is the initial endowment $\bar{M}$.

The critical points for maximizing the utility function $U = f(x,y)$, subject to budget constraint $\bar{M} = g(x,y)$, are found by employing the method of Lagrange multipliers. Consider the function

$$V(x,y,\lambda) = f(x,y) + \lambda \left[g(x,y) - \bar{M}\right],$$

where x, y, and λ are treated as variables. However, λ (the Lagrange multiplier) is actually a constant and not a variable. A necessary condition for maximization of $V(x,y,\lambda)$ requires that

$$\frac{\partial V}{\partial x} = 0; \qquad \frac{\partial V}{\partial y} = 0; \qquad \frac{\partial V}{\partial \lambda} = 0.$$

This system of equations constitutes the first-order conditions, and most students find this result easy to understand. Solving this system of equations yields the critical points, which correspond to a maximum, minimum, or saddle-point. However, many students find the second-order conditions difficult to conceptualize, as their derivation does not follow directly from $V(x,y,\lambda)$. This assertion is proven in Section 5.

In what follows, an alternative technique to Lagrange multipliers is used to derive both the first- and second-order conditions for maximization of the utility function,

$$U = f(x,y), \qquad \text{subject to budget constraint,} \qquad \bar{M} = g(x,y).$$

2. TRANSLATION INTO A FUNCTION OF ONE VARIABLE

Given the budget constraint $\bar{M} = g(x,y)$, let $x = h(t)$ and $y = k(t)$ be a parametric representation of the constraint. Thus, $\bar{M} = g[h(t),\, k(t)]$. As t is allowed to vary, the budget constraint is traced out in the plane. Applying the same transformation to the utility function $U = f(x,y)$, we have $U(t) = f[h(t),\, k(t)]$, where the point $(h(t),\, k(t))$ lies on the constraint. Notice that the utility function $U = f(x,y)$ is now expressed as a single-valued function of t, specifically,

$$U(t) = f[h(t), k(t)].$$

3. DERIVATION OF FIRST-ORDER CONDITIONS

Applying the necessary conditions for maximization of a single-value function yields

$$\frac{dU(t)}{dt} = \left(\frac{\partial f}{\partial x}\right)\left(\frac{dx}{dt}\right) + \left(\frac{\partial f}{\partial y}\right)\left(\frac{dy}{dt}\right) = 0. \tag{1}$$

For simplicity, let $\frac{\partial f}{\partial x} = f_x$; $\frac{\partial f}{\partial y} = f_y$; $\frac{dx}{dt} = \dot{x}$; $\frac{dy}{dt} = \dot{y}$.
Making the appropriate substitutions, equation (1) is restated as

$$\frac{dU(t)}{dt} = f_x\dot{x} + f_y\dot{y} = 0.$$

However, since $\bar{M}$ is a constant,

$$\frac{d\bar{M}}{dt} = g_x\dot{x} + g_y\dot{y} = 0.$$

Let $\nabla f = (f_x, f_y)$ and $\nabla g = (g_x, g_y)$ represent the gradient vectors of f and g, respectively. Let $\vec{T} = (\dot{x}, \dot{y})$ represent the tangent vector to $\bar{M} = g(x,y)$.

In vector notation, $f_x\dot{x} + f_y\dot{y} = 0$ can be written as $\nabla f \cdot \vec{T} = 0$, and $g_x\dot{x} + g_y\dot{y} = 0$ can be written as $\nabla g \cdot \vec{T} = 0$, where the operation • refers to the dot product. However, $\nabla f \cdot \vec{T} = 0$

implies that ∇f and $\vec{T}$ are perpendicular. Likewise, $\nabla g \cdot \vec{T} = 0$ implies that ∇g and $\vec{T}$ are perpendicular. Therefore, the two vectors ∇f and ∇g are perpendicular to the same vector $\vec{T}$.

If two vectors are perpendicular to the same vector in two dimensions, then, one vector is a scalar multiple of the other. In other words, ∇f and ∇g are parallel to each other. Therefore, $\nabla f + \lambda \nabla g = \underline{0}$, for some appropriate constant λ, which is the familiar Lagrange multiplier, where $\underline{0}$ is the zero-vector $(0,0)$. Thus, the necessary conditions for the above constrained maximization problem can be stated as follows:

- $f_x + \lambda g_x = 0$,
- $f_y + \lambda g_y = 0$,
- $\bar{M} = g(x,y)$.

PROOF. A well-known result is that

$$\nabla f + \lambda \nabla g = \underline{0} = (f_x, f_y) + \lambda (g_x, g_y) = (f_x + \lambda g_x, f_y + \lambda g_y) = \underline{0}.$$

Therefore, $f_x + \lambda g_x = 0$ and $f_y + \lambda g_y = 0$. Obviously, the critical point must lie on constraint $\bar{M} = g(x,y)$.

It is readily verifiable that applying the calculus to function $V(x,y,\lambda)$ yields identical necessary conditions for a maximum. However, as the next two sections demonstrate, this is not the case for the second-order conditions.

4. DERIVATION OF SECOND-ORDER CONDITIONS

For a true maximum of a single-valued function, it is required that $\frac{d^2U(t)}{dt^2} < 0$. Note that, since $\bar{M}$ is fixed as a function of t, $\frac{d\bar{M}}{dt} = 0$ and $\frac{d^2\bar{M}}{dt^2} = 0$.

Hence,

$$\frac{d\bar{M}}{dt} = g_x \dot{x} + g_y \dot{y} = 0.$$

Also,

$$\frac{d^2\bar{M}}{dt^2} = g_x \ddot{x} + g_y \ddot{y} + \left(\dot{x}^2 g_{xx} + 2\dot{x}\dot{y} g_{xy} + \dot{y}^2 g_{yy}\right) = 0.$$

Therefore,

$$\nabla g \cdot (\ddot{x} i + \ddot{y} j) + \left(\dot{x}^2 g_{xx} + 2\dot{x}\dot{y} g_{xy} + \dot{y}^2 g_{yy}\right) = 0, \qquad \text{where } i = (1,0) \quad \text{and} \quad j = (0,1), \tag{2}$$

and

$$\frac{d^2U(t)}{dt^2} = \nabla f \cdot (\ddot{x} i + \ddot{y} j) + \left(\dot{x}^2 f_{xx} + 2\dot{x}\dot{y} g_{xy} + \dot{y}^2 g_{yy}\right) < 0. \tag{3}$$

We know that $\nabla f + \lambda \nabla g = \underline{0}$. Multiply both sides of equation (2) by λ, so

$$\underline{0} = \lambda \nabla g \cdot (\ddot{x} i + \ddot{y} j) + \left(\dot{x}^2 \lambda g_{xx} + 2\dot{x}\dot{y} \lambda g_{xy} + \lambda \dot{y}^2 g_{yy}\right). \tag{4}$$

Adding equation (3) and equation (4) yields

$$(\nabla f + \lambda \nabla g) \cdot (\ddot{x} i + \ddot{y} j) + \dot{x}^2 (f_{xx} + \lambda g_{xx}) + 2\dot{x}\dot{y} (f_{xy} + \lambda g_{xy}) + \dot{y}^2 (f_{yy} + \lambda g_{yy}) < \underline{0}.$$

Note that $(\nabla f + \lambda \nabla g) = \underline{0}$. This result can be expressed as a quadratic form, which must be negative definite for a true maximum, subject to condition $(\dot{x}, \dot{y}) = k(-g_y, g_x)$, where k is a constant. Furthermore, note that $F = f + \lambda g$. Therefore,

$$(\dot{x}, \dot{y}) \begin{pmatrix} F_{xx} & F_{yx} \\ F_{xy} & F_{yy} \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} < 0. \tag{5}$$

Since $g[h(t),\, k(t)] = \bar{M}$, we know that $\frac{d\bar{M}}{dt} = g_x\dot{x} + g_y\dot{y} = 0$. The observation that

$$(\dot{x}, \dot{y}) \cdot (g_x, g_y) = 0,$$

implies that the two vectors are perpendicular. Therefore, $(\dot{x}, \dot{y})$ is parallel to

$$(-g_y, g_x) \quad \text{and} \quad (\dot{x}, \dot{y}) = \bar{k}(-g_y, g_x),$$

where $\bar{k}$ is some constant.

Making this substitution into equation (5) yields:

$$\bar{k}\,(-g_y, g_x)\begin{pmatrix} F_{xx} & F_{yx} \\ F_{xy} & F_{yy} \end{pmatrix}\bar{k}\begin{pmatrix} -g_y \\ g_x \end{pmatrix} < 0,$$

$$\bar{k}^2\left(F_{xx}g_x^2 - 2F_{xy}g_xg_y + F_{yy}g_y^2\right) < 0,$$

which implies that

$$\left(F_{xx}g_x^2 - 2F_{xy}g_xg_y + F_{yy}g_y^2\right) < 0. \tag{6}$$

The second-order conditions are represented by equation (6). Frequently, equation (6) is expressed in determinant form as

$$\begin{vmatrix} F_{xx} & F_{xy} & -g \\ F_{yx} & F_{yy} & -g_y \\ -g_x & -g_y & 0 \end{vmatrix} < 0.$$

This expression is known as the bordered Hessian determinant, named after its inventor, German mathematician Ludwig Otto Hesse (1811–1874).

5. DERIVATION OF SECOND-ORDER CONDITIONS FROM $V(x, y, \lambda)$

The sufficiency condition derived from $V(x, y, \lambda)$ is that the relevant quadratic form must be negative definite.

$$(\alpha, \beta, \gamma)\begin{pmatrix} F_{xx} & F_{yx} & g_x \\ F_{xy} & F_{yy} & g_y \\ g_x & g_y & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} < 0, \tag{7}$$

where $F = f + \lambda g$. Expanding equation (7), we obtain:

$$(\alpha, \beta)\begin{pmatrix} F_{xx} & F_{yx} \\ F_{xy} & F_{yy} \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \gamma\,(2\alpha g_x + 2\beta g_y) < 0.$$

The first expression is a quadratic form, but note that the second expression is linear in (g_x, g_y). Since γ can take on any value independent of α and β, we can find a suitable γ that makes the expression on the left-hand side of equation (7) > 0. This implies that we have a saddle-point and not a maximum.

Applying the calculus to function $V(x, y, \lambda)$, yields identical necessary conditions for a maximum, but the sufficiency condition derived from $V(x, y, \lambda)$ implies that the critical point is a saddle-point and not a true maximum. It is here where the equivalent relationship between $\max V(x, y, \lambda)$ and $\max f(x, y)$, subject to constraint $M = g(x, y)$, breaks down.

Postscripts

A classic in the mathematical treatment of advanced microeconomics is *Microeconomic Theory: A Mathematical Approach* by Henderson and Quandt [1], which has seen numerous editions since the first, in 1958. In particular, in the second edition, Henderson and Quandt [1, p. 404] state that maximizing $V(x, y, \lambda)$ is equivalent to maximizing $U = f(x, y)$, subject to budget constraint $\bar{M} = g(x, y)$. In subsequent editions, the authors omit this assertion, which is obviously incorrect. More recent treatment of the subject is found in Dixit [2], Varian [3], and Jehle and Reny [4], but in no case is rigorous derivation of the second-order conditions presented.

REFERENCES

1. J.M. Henderson and R.E. Quandt, *Microeconomic Theory: A Mathematical Approach*, McGraw-Hill, New York, (1971).
2. A.K. Dixit, *Optimization in Economic Theory*, Oxford University Press, Oxford, (1990).
3. H.R. Varian, *Microeconomic Analysis*, W.W. Norton & Company, New York, (1992).
4. G.A. Jehle and P.J. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, New York, (1997).